2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Peak Time-Windowed Mean Estimation using Convex Optimization

Jared Miller[1], Niklas Schmid[1], Matteo Tacchi[2], Didier Henrion[3], Roy S. Smith[1]

***Abstract*— This paper presents an algorithmic approach towards bounding the peak time-windowed average value attained by a state function along trajectories of a dynamical system. An example includes the maximum average current flowing across a power line in any 5-minute window. The peak time-windowed mean estimation task may be posed as a finite-dimensional but nonconvex optimization problem in terms of an initial condition and stopping time. This problem can be lifted into an infinite-dimensional linear program in occupation measures, where no conservatism is introduced under compactness and dynamical regularity assumptions. The peak time-windowed mean estimation linear program is in turn truncated into a convergent sequence of semidefinite programs using the moment-Sum-of-Squares hierarchy. Bounds of the time-windowed mean are computed for example systems.**

I. Introduction

Peak estimation is the practice of finding the maximal (or minimal) value that a state function p attains along trajectories of a dynamical system [1]. Examples of peak estimation problems (or bounding extreme events) include finding the maximum angular velocity of a motor, bounding the height of a wave, or determining the distance of closest approach between two autonomous agents (e.g. aircraft) [2].

In some settings, the instantaneous peak value attained by a trajectory may offer an incomplete understanding of the underlying safety properties. As a motivating example, a synchronous machine (motor) experiences a large current draw when first turning on, which will dissipate over the course of its transient response. This initial inrush current could exceed the rated capacity of the motor by a factor of 5-6 times without damaging motor operation [3]. The relevant quantity to measure safety would instead be the component's obedience to its thermal limits. These thermal limits can be associated with the time-windowed average current passing through the component, where the window's time horizon is related to the component's ability to dissipate heat.

Based on this motivation, we analyze the peak time-windowed mean-value of p along trajectories evolving according to Ordinary Differential Equation (ODE) dynamics within a state set $X \subset \mathbb{R}^n$. Trajectories can begin within an initial set $X_0 \subseteq X$ and will propagate for T time units. The peak time-windowed mean estimation problem, given a state function p, an ODE dynamics function f, and a time window $h \in [0, T]$, is posed as follows:

Problem 1.1: Find an initial condition x_0^* and a terminal time t^* to supremize

$$P^* = \sup_{t^*,\ x_0^*} \ \frac{1}{h} \int_{t'=t^*-h}^{t^*} p(x(t' \mid x_0^*)) dt' \tag{1a}$$

$$\text{s.t.} \quad \dot{x}(t) = f(t, x(t)) \quad \forall t \in [0, T] \tag{1b}$$

$$x(0) = x_0^* \tag{1c}$$

$$t^* \in [h, T],\ x_0^* \in X_0. \tag{1d}$$

In the case where $h = T$, Problem (1) can be interpreted as a Lagrange-type optimal control problem with a pure stage cost $\int_{t'=0}^{T} [p(x(t' \mid x_0^*))/T] dt'$ with $t^* = T$ [4], [5]. In the limit as $h \to 0$, Problem (1) will approach the instantaneous peak estimation problem (a free-terminal-time terminal cost $p(x(t^*))$ and a zero stage cost) when p is continuous [1].

Figures 1 and 2 illustrate the difference between the instantaneous and time-averaged peak. The signal $p(t)$ is plotted in black over a time horizon of $T = 8$. The red dotted lines plot the time-windowed average with $h = 1.5$. The blue square at $t_i = 2.0017$ is the instantaneous peak of $p_i = 1.5435$. The time-windowed average at t_i is $\int_{t'=t_i-1.5}^{t_i} p(t')dt' = 0.0173$. The blue star at $t_w = 5.7502$ achieves the maximal time-windowed average peak of $p_w = 0.9379$, while the instantaneous value of $p(t_a)$ is 0.8220. The magenta curve of Figure 1 is the region that is averaged to produce p_w.

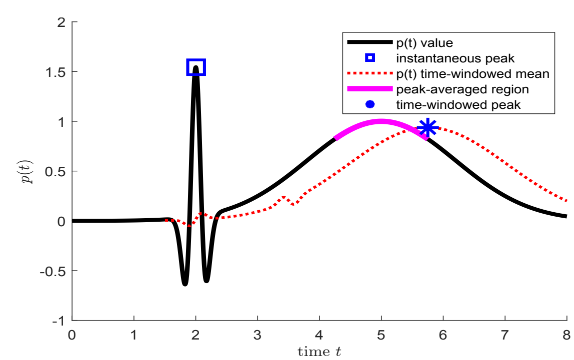

Fig. 1: Comparison of instantaneous peak (square) and time-windowed average peak (star) of a signal $p(t)$ (black curve)

Problem 1.1 is a finite-dimensional optimization problem in (t^*, x_0^*) but is generically nonconvex. We will lift the formulation in (1) into an infinite-dimensional Linear Program (LP) in occupation measures. Instances of occupation measure LP relaxations used for dynamical systems problems

[1]J. Miller, N. Schmid, and R. S. Smith are with the Automatic Control Laboratory (IfA) and NCCR Automation, Department of Information Technology and Electrical Engineering (D-ITET), ETH Zürich, Physikstrasse 3, 8092, Zürich, Switzerland ({jarmiller,nikschmid, rsmith}@control.ee.ethz.ch).

[2] M. Tacchi is with Univ. Grenoble Alpes, CNRS, Grenoble INP (Institute of Engineering Univ. Grenoble Alpes), GIPSA-lab, 38000 Grenoble, France. (e-mail: matteo.tacchi@gipsa-lab.fr)

[3] D. Henrion is with LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France; and the Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic. (henrion@laas.fr)

J. Miller and R. S. Smith are partially supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

N. Schmid is supported by the European Research Council under the Horizon 2020 Advanced under Grant 787845 (OCAL).

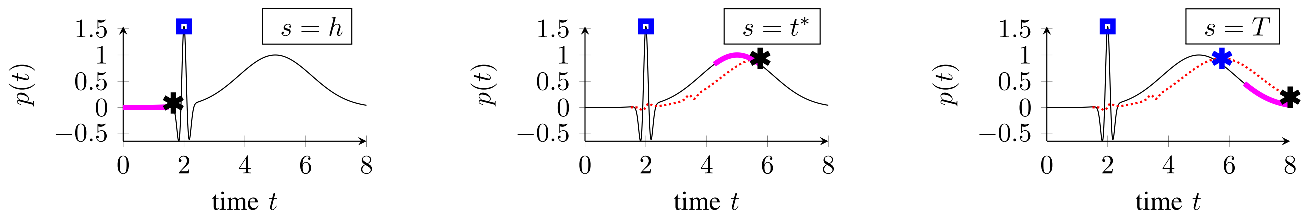

Fig. 2: Comparison of instantaneous peak (square) to time-windowed average peak (blue star) of a signal $p(t)$ from Figure 1 with stopping time $s \in \{h,\ t^*,\ T\}$. The magenta region highlights $p(t)$ in times $[s-h, s]$. The time-windowed average in this range is marked with the black star. The red dotted line depicts all time-windowed averages up to $t \leq s$.

includes optimal control [4], instantaneous peak estimation [1], [6], [7], reachable set estimation [8], and (stochastic) risk estimation [9]. To the best of our knowledge, the peak time-windowed mean estimation problem for ODEs has not been previously considered in the literature.

The infinite-dimensional LP must then be truncated into a finite-dimensional optimization problem in order to admit computational solutions. In this work, we will use the moment-Sum of Squares (SOS) hierarchy to form this discretization [10], yielding a sequence of Semidefinite Programs (SDPs) in increasing complexity whose solution values will converge to the optimal value P^* (under compactness, regularity, and polynomial structure conditions). Alternative methods include gridding (e.g. discrete-time-and-state Markov Decision Processes) [1], random sampling [11], and neural network verification [12].

This paper has the following structure: Section II provides an overview of preliminaries and notation. Section III proposes an infinite-dimensional LP in occupation measures that has the same optimal value as the peak time-windowed average cost problem (under compactness and regularity assumptions). Section IV truncates the LP into a convergent sequence of SDP using the moment-SOS hierarchy. Section V demonstrates this approach on an example system. Section VI concludes the paper. An extended version of this paper including stochastic processes, proofs of duality, and non-mean risk measures is available at [13].

II. Preliminaries

A. Notation and Measure Theory

The n-dimensional real Euclidean vector space is $\mathbb{R}^n$. The set of natural numbers is $\mathbb{N}$, and the set of k-dimensional multi-indices is $\mathbb{N}^k$. The set of natural numbers between a and b is $a:b \subset \mathbb{N}$. For a vector $x \in \mathbb{R}^n$ and a multi-index $\alpha \in \mathbb{N}^n$, exponentiation will be notated by $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$. The symbol $\mathbb{E}$ will denote the expectation of a random variable, and $\mathbb{P}$ will denote the probability of an event.

Let $K \in \mathbb{R}^n$ be a compact set and consider the vector space $C(K)$ of continuous functions from K to $\mathbb{R}$. The norm $\|f\|_{C(K)} := \max\{|f(x)| \mid x \in K\}$ equips $C(K)$ with a Banach structure [14]. In particular, one can define the dual space $C(K)'$ as the space of continuous linear functionals from $C(K)$ to $\mathbb{R}$ with duality notation $\langle f, \mu \rangle = \int f \, d\mu \in \mathbb{R}$ for $f \in C(K)$ and $\mu \in C(K)'$. By defining the nonnegative subcone $C_+(K) = \{f \in C(K) \mid \forall x \in K,\ f(x) \geq 0\}$, we can introduce the Riesz-Markov characterization of Radon (nonnegative) measures:

$$\mathcal{M}_+(K) = \{\mu \in C(K)' \mid \forall f \in C_+(K),\ \langle f, \mu \rangle \geq 0\}.$$

In the rest of this work, we adopt the following standard convention for (partial) derivatives of Radon measures, defined by their action on continuously differentiable functions as $\langle v, \partial\mu/\partial x_i \rangle = -\langle \partial v/\partial x_i, \mu \rangle$.

The mass of a nonnegative measure $\mu \in \mathcal{M}_+(K)$ is $\mu(K) = \langle 1, \mu \rangle$. The measure μ is a *probability measure* if its mass is 1. Given a point $x' \in K$, the Dirac delta $\delta_{x=x'} \in \mathcal{M}_+(K)$ is the unique probability measure supported only at $x = x'$ ($\forall f \in C(K):\ \langle f, \delta_{x=x'} \rangle = f(x')$). The pushforward of μ through a map $q: K \to L$ is written as as $q_\# \mu \in \mathcal{M}_+(L)$is the unique measure satisfying $\forall g \in C(L):$ $\langle g, q_\# \mu \rangle = \langle g \circ q, \mu \rangle$. Given measures $\mu \in \mathcal{M}_+(K)$ and $\nu \in \mathcal{M}_+(L)$ with support sets K, L, the product measure $\mu \otimes \nu \in \mathcal{M}_+(K \times L)$ is the unique measure satisfying $\forall f \in C(K), g \in C(L):\ \langle f \cdot g, \mu \otimes \nu \rangle = \langle f, \mu \rangle \cdot \langle g, \nu \rangle$, where $f \cdot g = K \times L \ni (x, y) \longmapsto f(x) \cdot g(y)$.

B. Occupation Measures

Consider a dynamical system $x(t) \in \mathbb{R}^n$, $t \geq 0$. Notice that, if the initial condition $x(0)$ is set to be a random variable, or if the dynamics depend on a random control policy, then the state $x(t)$ will be a random variable at all times $t > 0$. In such case, the trajectories and their probability laws can be encoded over a finite time horizon $T > 0$ into a single Radon measure called the *occupation measure*. The occupation measure is defined as follows: set $K = [0, T] \times X$ where X is a compact observation set with $\mathbb{P}(x(0) \in X) = 1$, as well as a stopping time

$$s \leq \sup\{t \in [0, T] \mid \mathbb{P}(x(t) \in X) = 1\}.$$

The occupation measure $\mu \in \mathcal{M}_+(K)$ is defined through its duality with functions $v \in C(K)$ as $\langle v, \mu \rangle = \mathbb{E}\left[\int_0^s v(t, x(t))\, dt\right]$, i.e. the expected accumulated value of v along observed trajectories. Given the probability law μ_0 of the initial condition (which could include a Dirac measure $\mu_0 = \delta_{x=x_0'}$), the occupation measure is solution to a Kolmogorov-type PDE of the form

$$\mu_s - \mathcal{L}^\dagger \mu = \delta_0 \otimes \mu_0 \tag{2}$$

where $\mu_s \in \mathcal{M}_+(K)$ is an unknown stopping measure encoding a distribution of time-states $(s, x(s))$, and the operator $\mathcal{L}^\dagger$ encodes the system's dynamics. In the case of an uncontrolled ordinary differential equation $\dot{x}(t) = f(t, x(t))$, the operator $\mathcal{L}^\dagger$ is obtained by testing (2) against test function v to derive

$$\begin{aligned}\langle v, \mathcal{L}^\dagger \mu\rangle &= \langle v, \mu_s - \delta_0 \otimes \mu_0\rangle \\ &= \mathbb{E}_{x(0)\sim\mu_0}\left[v(s, x(s)) - v(0, x(0))\right] \\ &= \mathbb{E}_{x(0)\sim\mu_0}\left[\int_0^s \frac{d}{dt} v(t, x(t))\, dt\right] = \langle \mathcal{L}v, \mu\rangle.\end{aligned}$$

Applying the chain rule $\mathcal{L}v = \partial v/\partial t + f^\top \operatorname{grad} v$ yields$\mathcal{L}^\dagger \mu = -\partial \mu/\partial t - \operatorname{div}(f\,\mu)$. In the ODE case, equation (2) is called the Liouville equation, and $\mathcal{L}$ is called the generator of the dynamics f. Any tuple (μ_0, μ, μ_s) satisfying (2) is called a *relaxed occupation measure*.

III. Time-Windowed Mean Linear Program

This section will relax the peak time-windowed average program in (1) into an infinite-dimensional LP.

A. Assumptions

We introduce the following assumptions:

A1 $0 < T < \infty$ and sets X_0, X are compact.
A2 The objective function $p(x)$ is continuous.
A3 The function $x \mapsto f(\cdot, x)$ is Lipschitz within $[0, T] \times X$.
A4 Trajectories starting from X_0 obey a non-return criterion: if $\exists (t^*, x_0^*) \in [0, T] \times X_0$ such that $x(t^* \mid x_0^*) \notin X$, then $\forall t \geq t^* :\ x(t \mid x_0^*) \notin X$.

Technicalities of the non-return assumption A4 are explained further in Remark 1 of [2]. Non-return implies that all stopped trajectories maximizing the time-windowed valuation of p will stay in X, and therefore can be tracked by an occupation measure μ supported in X.

B. Augmented Time

To phrase linear programs for (1), we will first define a new constant state s to the dynamics ($\dot{s} = 0$). The constant state s will serve as the stopping time t^* in optimization problem (1). Since at the stopping time it holds that $t = t^\star = s$, the stopping measure μ_s of Section II-B can therefore be decomposed as $\mu_s = \varphi_\# \mu_\tau$ for some terminal measure $\mu_\tau \in \mathcal{M}_+([h, T] \times X)$ pushed forward through the map $\varphi = (s, x) \longmapsto (s, s, x)$ i.e. for $v \in C([h, T]^2 \times X)$ it holds

$$\langle v, \mu_s\rangle = \langle v|_{t=s}, \mu_\tau\rangle = \int v(s, s, x)\, d\mu_\tau(s, x). \quad (3)$$

We thereby define the augmented-time support sets as

$$\Omega_+ = \{(s, t) \in [h, T] \times [0, T] \mid t \in [s - h, s]\} \quad (4a)$$
$$\Omega_- = \{(s, t) \in [h, T] \times [0, T] \mid t \in [0, s - h]\}. \quad (4b)$$

C. Infinite dimensional LP

A mean-type time-windowed risk LP from (1) will be formulated in terms of measures $(\mu_0, \mu_\tau, \mu_+, \mu_-)$:

$$\begin{aligned}
&\mu_0(s, x) \in \mathcal{M}_+([h, T] \times X_0) && \text{Initial} && (5a)\\
&\mu_\tau(s, x) \in \mathcal{M}_+([h, T] \times X) && \text{Terminal} && (5b)\\
&\mu_+(s, t, x) \in \mathcal{M}_+(\Omega_+ \times X) && \text{Risk Occ.} && (5c)\\
&\mu_-(s, t, x) \in \mathcal{M}_+(\Omega_- \times X) && \text{Past Occ.} && (5d)
\end{aligned}$$

The mean-type measure program for (1) is:

Problem 3.1: Find an initial measure μ_0, a terminal measure μ_τ, a risk occupation measure μ_+, and a past occupation measure μ_- to supremize:

$$\begin{aligned}
p^* = \ &\sup \quad \langle p, \mu_+\rangle / h && (6a)\\
&\text{s.t. } \varphi_\# \mu_\tau = \delta_{t=0} \otimes \mu_0 + \mathcal{L}^\dagger(\mu_- + \mu_+) && (6b)\\
&\langle 1, \mu_0\rangle = 1 && (6c)\\
&\langle 1, \mu_+\rangle = h && (6d)\\
&\text{Support constraints in (5).} && (6e)
\end{aligned}$$

The sum $\mu_- + \mu_+$ serves as the relaxed occupation measure for dynamics with generator $\mathcal{L}$ as in Section II-B. The mean value of p is only evaluated in the range $t \in [s - h, s]$, which is enforced by the Ω_+ support constraint in (5b). Constraint (6d) ensures that trajectories defined in μ_+ will be recorded for exactly h time units, ensuring that these trajectories are well-defined when taking the time-windowed average operation.

Lemma 3.2: Under only Assumption A4, program (6) is an upper-bound on (1) ($p^* \geq P^*$).

Proof: This proof will construct feasible measures in (5) from a feasible point $(t^*, x_0^*) \in [h, T] \times X_0$ of (1). The initial measure μ_0 may be set to $\mu_0 = \delta_{s=t^*,\ x=x_0}$. Letting $x(t^*)$ represent the probability distribution of the process at time t^*, the terminal measure may be chosen as $\mu_\tau = \delta_{s=t^*,\ x=x(t^*)}$. Defining $\mu(t, x)$ as the occupation measure of the process $x(t)$ between times 0 and t^*, let $\mu_{[0, t^*-h]}(t, x)$ and $\mu_{[t^*-h, t]}(t, x)$ refer to the restrictions of μ in times $[0, t^* - h]$ and $[t^* - h, t^*]$ respectively. The decomposed relaxed occupation measures may therefore be chosen as $\mu_- = \delta_{s=t^*} \otimes \mu_{[0, t^*-h]}$ and $\mu_+ = \delta_{s=t^*} \otimes \mu_{[t^*-h, t^*]}$. Note that constraint (6d) is satisfied because μ_+ is defined over exactly h time units.

Because there exists an injective map between feasible point (t^*, x_0^*) from (1) and feasible measures (5) for (6b)-(6e) such that the cost is preserved, the optimal value p^* for (6) must therefore be an upper-bound for (1) as $p^* \geq P^*$. ∎

The main result of our paper is the following theorem:

Theorem 3.3: Under Assumptions A1-A4, there is no relaxation gap ($p^* = P^*$).

Proof: Let (μ_0, μ_τ, μ_J) (with $\mu_J \in \mathcal{M}_+([h, T] \times [0, T] \times X)$) satisfy

$$\varphi_\# \mu_\tau = \delta_{t=0} \otimes \mu_0 + \mathcal{L}^\dagger \mu_J. \quad (7)$$

By Assumptions A1-A4 and Theorem 3.3 of [1], every feasible tuple (μ_0, μ_τ, μ_J) satisfying (6b)-(6e) is supported on the graph of stochastic trajectories of $\mathcal{L}$.

As such, there exists a decomposition of μ_J into μ_- and μ_+ through the (s, t) temporal supports in (4) subject to the requirement (6d). Such a decomposition may be achieved through the restriction procedure from Theorem 3.1 of [15], in which μ_+ is chosen to maximize $\langle t, \mu_+\rangle$ such that $\mu_+ + \mu_- = \mu_J$, $\langle 1, \mu_+\rangle = h$ under (5c)-(5d). It therefore holds that the tuple $(\mu_0, \mu_\tau, \mu_- + \mu_+)$ is supported on the graph of a stochastic process, and all measures satisfy the support constraints in (5). Specifically, extremal trajectories can be found among Dirac measures that correspond to single initial conditions in Problem 1.1. Given that μ_0 is a probability distribution by (6c), the term $\langle p, \mu_+\rangle/h$ in the objective (6a) evaluates to the integral in (1a). The absence of a relaxation gap is therefore proven under A1-A4. ■

IV. Time-Windowed Average Finite Truncation

The infinite-dimensional LPs presented in Section III must be truncated into finite-dimensional programs in order to be computationally tractable. This section will first present the moment-SOS hierarchy of SDPs for truncation of measure LP, and will then apply this hierarchy towards grid-free truncation of (6) (see [10] for more details).

A. Moment-SOS Background

Problem (6) involves infinite-dimensional decision variables (namely Radon measures) as well as infinite-dimensional linear constraints of the form $\mathcal{A}\,\mu = b$ such as the Kolmogorov equation (6b). To cope with such infinite dimension, the moment-SOS framework consists of three steps: *(i)* testing infinite dimensional constraints against polynomials, *(ii)* representing Borel measures $\mu \in \mathcal{M}_+(X)$ with their moments $y_\alpha = \langle x^\alpha, \mu\rangle$ and *(iii)* truncating the degree of both moment variables and constraints to a finite value $k < \infty$. The sufficiency of testing over polynomials in step *(i)* is ensured through the Stone-Weierstraß theorem. Step *(ii)* is made possible by Putinar's Positivstellensätz [16, Lemma 3], which characterizes moment sequences among multi-indexed real sequences:

Lemma 4.1 (Putinar's Positivestellensätz): [16]] Let $g_1, \ldots, g_m \in \mathbb{R}[x]$ define a Basic Semialgebraic (BSA) set

$$X = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}.$$

Let $y = (y_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ be a multi-indexed real sequence with associated Riesz linear functional

$$L_y = \left\{ \begin{array}{ccc} \mathbb{R}[x] & \longrightarrow & \mathbb{R} \\ \sum_\alpha p_\alpha\, x^\alpha & \longmapsto & \sum_\alpha p_\alpha\, y_\alpha. \end{array} \right.$$

Then, assuming the Archimedean property for basic semi-algebraic sets ($g_m(x) = R^2 - x^\top x$, which is equivalent to compactness of X up to adding a redundant ball constraint to its description), there exists a Radon measure $\mu \in \mathcal{M}_+(X)$ such that $\forall \alpha \in \mathbb{N}^n, \quad y_\alpha = \langle x^\alpha, \mu\rangle$ if and only if the following quadratic forms are nonnegative:

$$Q_y = p \mapsto L_y(p^2) \geq 0, \quad Q_{g_i y} = p \mapsto L_y(g_i\, p^2) \geq 0. \quad (8)$$

Notice that when the conditions of Lemma 4.1 are met, it holds that $\forall p \in \mathbb{R}[x], \quad L_y(p) = \langle p, \mu\rangle$. Hence, Lemma 4.1 means that it is possible to replace Radon measures in (6) with real sequences y complemented with appropriate positivity constraints on $(Q_y, Q_{g,y})$.

Step *(iii)* truncates infinite dimensional moment sequences to finite size moment vectors. For a degree $k \in \mathbb{N}$ and a polynomial $g \in \mathbb{R}[x]$, let $M_k(gy)$ be the symmetric square matrix of size $\binom{n+k-\lceil \deg g/2\rceil}{n}$ defined by

$$M_k(gy)_{\alpha,\beta} = L_y(g(x)\, x^{\alpha+\beta}) = \textstyle\sum_\gamma g_\gamma y_{\alpha+\beta+\gamma}. \quad (9)$$

The finite-dimensional matrix $M_k(gy)$ is equal to the top corner of the inifinite-dimensional matrix representing the quadratic form Q_{gy} in the basis of monomials. The degree-k truncation of a measure LP in step (iii) involves replacing each measure variable by a sequence of (pseudo)-moments y up to degree $2k$, imposing that $M_k(y)$ and $M_k(g_i y)$ are each Positive Semidefinite (PSD) matrices, and replacing linear constraints in measures $\mathcal{A}\mu = b$ with a finite number of constraints in the pseudo-moments (e.g. $L_y(\langle x^\alpha, b - \mathcal{A}\,\mu\rangle) = 0, \ \forall |\alpha| \leq k$). Each degree-$k$ truncation may be solved by SDP algorithms, and the process of increasing $k \to \infty$ to achieve better bounds is the moment-SOS hierarchy.

B. Peak Time-Windowed Mean Moment-SOS Programs

In order to apply the Moment-SOS hierarchy, assumptions of polynomial structure must be imposed:

- A5 Both $p(x)$ and $f(t, x)$ are polynomial maps.
- A6 The sets X_0, X each have an Archimedean BSA representation:

$$X = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$
$$X_0 = \{x \in \mathbb{R}^n \mid g_{01}(x) \geq 0, \ldots, g_{0\ell}(x) \geq 0\}.$$

To each degree $k \in \mathbb{N}$ and function f with generator $\mathcal{L}_f$, the associated dynamics degree $\tilde{k} \geq k$ can be defined as

$$\tilde{k} = \left\lceil \max_{v\in\mathbb{R}_{2k}[t,x]} \frac{\deg(\mathcal{L}_f v)}{2} \right\rceil = k + \left\lceil \frac{\deg(f) - 1}{2} \right\rceil. \quad (10)$$

In addition, for $y = (y_\alpha)_{|\alpha|\leq 2k} \in \mathbb{N}^n_{2k}$ and $g \in \mathbb{R}_{2k}[x]$, let $\hat{k} = k - \lceil \deg(g)/2 \rceil$ and define the localizing matrix $M_k(g\,y)$ as the matrix representation of Q_{gy} in a basis of $\mathbb{R}_{\hat{k}}[x]$. We define the following polynomials:

$$g_h(s) = (T - s)(s - h)$$
$$g_+(s, t) = (s - t)(t - s + h)$$
$$g_-(s, t) = (s - t - h)t$$

so that it holds

$$[h, T] = \{s \in \mathbb{R} \mid g_h(s) \geq 0\}$$
$$\Omega_+ = \{(s, t) \in \mathbb{R}^2 \mid g_h(s) \geq 0, g_+(s, t) \geq 0\}$$
$$\Omega_- = \{(s, t) \in \mathbb{R}^2 \mid g_h(s) \geq 0, g_-(s, t) \geq 0\}.$$

The degree-k moment truncation of (6) from Problem 3.1 will be posed in terms by forming finite-vector pseudomoment sequences y as $(\mu_0, \mu_\tau, \mu_+, \mu_-) \to (y^0, y^\tau, y^+, y^-)$. The restriction of the Liouville relation in (6b) as parameterized by $\alpha \in \mathbb{N}, \ \beta \in \mathbb{N}, \gamma \in \mathbb{N}^n$ is:

$$\begin{aligned} \mathrm{Liou}_{\alpha\beta\gamma} = {} & L_{y^\tau}(t^{\alpha+\beta} x^\gamma) - L_{y^0}(s^\alpha \delta_{\beta=0} x^\gamma) \\ & - L_{y^+ + y^-}(\hat{\mathcal{L}}(s^\alpha t^\beta x^\gamma)). \end{aligned} \quad (11)$$

The degree-k truncation of Problem 3.1 is

Problem 4.2: Find pseudo-moment sequences $y^0, y^\tau \in \mathbb{R}^{\mathbb{N}^{n+1}_{2k}}$, $y^+, y^- \in \mathbb{R}^{\mathbb{N}^{n+2}_{2k}}$ to maximize

$$p_k^\star = \max_{y^0, y^\tau, y^+, y^-} L_{y^+}(p)/h \tag{12a}$$

$$\text{s.t. } \mathrm{Liou}_{\alpha\beta\gamma}(y^0, y^\tau, y^+, y^-) = 0 \tag{12b}$$

$$\forall(\alpha, \beta, \gamma) \in \mathbb{N}^n$$

$$L_{y^0}(1) = y^0_0 = 1 \tag{12c}$$

$$L_{y^+}(1) = y^+_0 = h \tag{12d}$$

$$M_k(1\, y^\bullet) \succeq 0, \quad M_k(g_h y^\bullet) \succeq 0, \tag{12e}$$

$$\forall \bullet \in \{0, \tau, +, -\}$$

$$M_k(g_- y^-) \succeq 0, \quad M_k(g_+ y^+) \succeq 0. \tag{12f}$$

The degree-truncated Liouville relation in (12b) is defined by $\binom{n+2+2\tilde{k}}{2\tilde{k}} < \infty$ scalar linear equality constraints. Constraints (12c) and (12d) are scalar linear equality constraints, and the Linear Matrix Inequality (LMI) constraints (12e), (12f) have size at most $\binom{n+2+\tilde{k}}{\tilde{k}} < \infty$. The finite-dimensional Problem 4.2 may therefore be solved by SDP optimization methods, such as interior-point programs.

Theorem 4.3: Under Assumptions A1-A6, the objectives of (12) will satisfy $p_k^* \geq p_{k+1}^* \geq p_{k+2}^* \ldots$ and will converge to the optimal solution of Problem 3.1 as $\lim_{k\to\infty} p_k^* = p^*$.

Proof: See Appendix C of [13]. ■

C. Computational Complexity

The per-iteration complexity of solving an SDP derived from the degree-k Moment-SOS hierarchy in n variables is $O(n^{6k})$ and $O(k^{4n})$ [10]. This scaling is due to the $\binom{n+k}{k}$ size of PSD localizing matrix constraints. Table I reports the size of the maximal-size PSD matrix constraints (for each pseudo-moment sequence) from Problem 4.2. The scaling of solving Problem 4.2 by interior-point methods will therefore grow in a jointly polynomial manner as $(n+2)^{6\tilde{k}}$ or $\tilde{k}^{4(n+2)}$.

TABLE I: Size of PSD matrices needed to represent formulation (12) of Problem 4.2 at degree k

Measure	μ_0	μ_τ	μ_+	μ_-
Constraint	$M_k(y^0)$	$M_k(y^\tau)$	$M_{\tilde{k}}(y^+)$	$M_{\tilde{k}}(y^-)$
PSD Size	$\binom{n+1+k}{k}$	$\binom{n+1+k}{k}$	$\binom{n+2+\tilde{k}}{\tilde{k}}$	$\binom{n+2+\tilde{k}}{\tilde{k}}$

V. Numerical Example

MATLAB (2023b) code to generate all examples is publicly available[1]. Dependencies include Gloptipoly [17], YALMIP [18], and Mosek [19].

A. Two-Dimensional Flow System

The Flow system from [20]:

$$\dot{x} = \begin{bmatrix} x_2 \\ -x_1 - x_2 + \frac{1}{3}x_1^3 \end{bmatrix}. \tag{13}$$

This example involves time-windowed peak mean estimation of the function $p(x) = -x_2$ with an evaluated time window of $h = 1$ and a time horizon of $T = 5$. The considered state set is $X = [-3, 3]^2$, and the set of initial conditions is the disc $X_0 = \{x \in \mathbb{R}^2 \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2\}$. Solving Problem 4.2 results in the upper bounds $p^*_{1:5} = \{2.4192, 0.9094, 0.5667, 0.5628, 0.5624\}$. The peak instantaneous value of $p(x)$ is upper-bounded by 0.5734 (solving the peak estimation routine from [6] at $k = 4$).

Figure 3 visualizes trajectories (cyan) of system (13) starting from the black-circle initial set X_0. The dark blue curve is a trajectory extracted from moments of the $k = 4$ solution, with a maximal (sampled) time-windowed $p(x)$ value of 0.5600. A feasible trajectory extracted from the moment matrix solution (second largest eigenvalue $\leq 10^{-3}$) starts at $x_0^* \approx [1.4900, -0.3940]$ (blue circle) and reaches a time-windowed extremum at $x_p^* \approx [0.4125, -0.5371]$ (blue star) at time $t_p^* \approx 2.1537$. The window for averaging over $p(x)$ begins at time $t_p^* - h = 1.1537$ time units, which yields a spatial location of $x(t_p^* - h \mid x_0^*) \approx [0.9683, -0.5351]$ (blue triangle). The magenta region between the blue triangle and the blue star is the portion of the trajectory that is time-averaged (times 1.1537 to 2.1537).

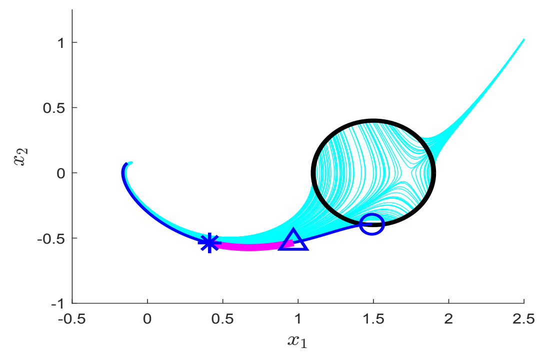

Fig. 3: Trajectories of (13) starting from the black-circle initial set X_0, with a near-optimal region highlighted for $p(x) = -x_2,\ h = 1$.

Figure 4 displays (in cyan) evaluations of the time-averaged costs $(1/h)\int_{t'=t-h}^{t} p(x(t' \mid x_0))dt'$ along trajectories from Figure 3 (where $x_0 \in X_0$ is the randomly sampled initial condition). The blue curve is the time-windowed average value of p along the near-optimal trajectory from Figure 3, which achieves its largest value at the blue star (time $t_p^* \approx 2.1537$). The red dotted line above all sampled objective curves is the $d_4^* = 0.5628$ bound.

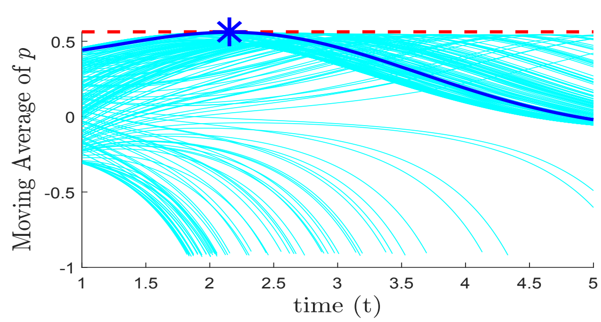

Fig. 4: Evaluation of time-windowed mean of $p(x) = -x_2,\ h = 1$ along trajectories in Figure 3.

[1] https://doi.org/10.3929/ethz-b-000662948

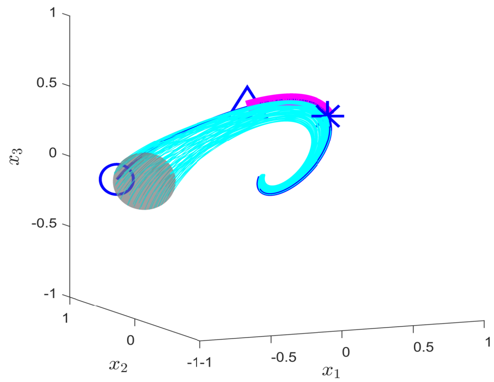

Fig. 5: Trajectories of Twist system (14) starting from the gray-sphere initial set X_0.

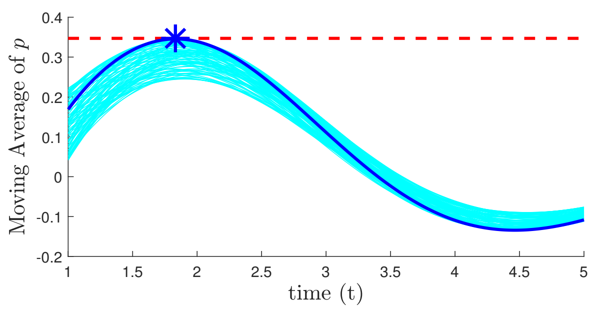

Fig. 6: Evolution of time-windowed mean of $p(x) = x_3, h = 1$ along trajectories in Figure 3.

B. Three-Dimensional Twist System

The Twist system from [2] is

$$\dot{x}_i(t) = \textstyle\sum_j A_{ij}x_j - B_{ij}(4x_j^3 - 3x_j)/2, \qquad (14)$$

$$A = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \qquad (15)$$

This example involves finding bounding the time-averaged value of $p(x) = x_3$ for parameters of $X = [-1, 1]^3$, $h = 1$, $T = 5$. Trajectories start in the initial set $X_0 = \{x \in \mathbb{R}^3 \mid (x_1 + 0.75)^2 + (x_2 - 0.4)^2 + (x_3 + 0.1)^2 \leq 0.2^2\}$. Figure 5 plots trajectories of (14) in cyan using the same coloring convention as in Figure 3. Solving Problem 4.2 at degrees $k \in 1..4$ results in the bounds $p^*_{1:4} = [0.4687, 0.3497, 0.3475, 0.3470]$. At the degree $k = 4$ relaxation, the approximately-optimal parameters of $x_0^* \approx [-0.8705, 0.5586, -0.1027]$, $x_p^* \approx [0.4593, 0.2197, 0.3043]$, and $t_p^* \approx [1.8315]$ are recovered. Figure 6 plots the time-windowed values of p along trajectories.

VI. Conclusion

This paper focuses on bounding the time-windowed average value of a state function $p(x)$ (Problem 1.1), rather than the previously considered instantaneous value of $p(x)$ [1], [6]. The peak time-windowed average estimation problem is converted into an infinite-dimensional LP in measures (Problem 4.2), which is in turn truncated into a sequence of SDPs using the moment-SOS hierarchy (Problem 4.2). There is no relaxation gap between the LP and the original problem under A1-A4 (Theorem 3.3), and under the further imposition of A5-A6, the sequence of SDPs will converge in objective to the true peak time-windowed average (Theorem 4.3). The key insight was in adding an extra time variable s subject to the augmented-temporal support constraints in (4). Future work will involve conditions for peak-minimizing control of the time-windowed averaged (risk) value of p.

VII. Acknowledgements

The authors thank Adrian Hauswirth for discussions about time-windowed risk estimation in power systems.

References

[1] M. J. Cho and R. H. Stockbridge, "Linear Programming Formulation for Optimal Stopping Problems," SIAM J. Control Optim., vol. 40, no. 6, pp. 1965–1982, 2002.
[2] J. Miller and M. Sznaier, "Bounding the Distance to Unsafe Sets with Convex Optimization," IEEE Trans. Autom., pp. 1–15, 2023.
[3] R. A. Turner and K. S. Smith, "Transformer inrush currents," IEEE Industry Applications Magazine, vol. 16, no. 5, pp. 14–19, 2010.
[4] R. Lewis and R. Vinter, "Relaxation of Optimal Control Problems to Equivalent Convex Programs," Journal of Mathematical Analysis and Applications, vol. 74, no. 2, pp. 475–493, 1980.
[5] D. Liberzon, Calculus of Variations and Optimal Control Theory: A Concise Introduction. Princeton University Press, 2011.
[6] G. Fantuzzi and D. Goluskin, "Bounding Extreme Events in Nonlinear Dynamics Using Convex Optimization," SIAM Journal on Applied Dynamical Systems, vol. 19, no. 3, pp. 1823–1864, 2020.
[7] J. Miller, D. Henrion, and M. Sznaier, "Peak Estimation Recovery and Safety Analysis," IEEE Control Systems Letters, vol. 5, no. 6, pp. 1982–1987, 2021.
[8] M. Korda, D. Henrion, and C. N. Jones, "Controller design and value function approximation for nonlinear dynamical systems," Automatica, vol. 67, pp. 54–66, 2016.
[9] J. Miller, M. Tacchi, M. Sznaier, and A. Jasour, "Peak Value-at-Risk Estimation for Stochastic Processes using Occupation Measures," 2023, arXiv:2303.16064.
[10] J. B. Lasserre, Moments, Positive Polynomials And Their Applications, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.
[11] P. Mohajerin Esfahani, T. Sutter, D. Kuhn, and J. Lygeros, "From infinite to finite programs: Explicit error bounds with applications to approximate dynamic programming," SIAM J. Optim., vol. 28, no. 3, pp. 1968–1998, 2018.
[12] A. Abate, D. Ahmed, A. Edwards, M. Giacobbe, and A. Peruffo, "FOSSIL: A Software Tool for the Formal Synthesis of Lyapunov Functions and Barrier Certificates using Neural Networks," in Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control, 2021, pp. 1–11.
[13] J. Miller, N. Schmid, M. Tacchi, D. Henrion, and R. S. Smith, "Peak time-windowed risk estimation of stochastic processes," 2024. [Online]. Available: https://arxiv.org/abs/2404.06961
[14] T. Tao, An Introduction to Measure Theory. American Mathematical Society Providence, RI, 2011, vol. 126.
[15] D. Henrion, J. B. Lasserre, and C. Savorgnan, "Approximate Volume and Integration for Basic Semialgebraic Sets," SIAM review, vol. 51, no. 4, pp. 722–743, 2009.
[16] M. Putinar, "Positive Polynomials on Compact Semi-algebraic Sets," Indiana University Mathematics Journal, vol. 42, no. 3, pp. 969–984, 1993.
[17] D. Henrion and J.-B. Lasserre, "GloptiPoly: Global Optimization over Polynomials with Matlab and SeDuMi," ACM Transactions on Mathematical Software (TOMS), vol. 29, no. 2, pp. 165–194, 2003.
[18] J. Lofberg, "YALMIP : a toolbox for modeling and optimization in MATLAB," in ICRA (IEEE Cat. No.04CH37508), 2004, pp. 284–289.
[19] M. ApS, The MOSEK optimization toolbox for MATLAB manual. Version 10.1., 2023.
[20] S. Prajna and A. Jadbabaie, "Safety Verification of Hybrid Systems Using Barrier Certificates," in International Workshop on Hybrid Systems: Computation and Control. Springer, 2004, pp. 477–492.